Mapping Tabular Data
Problem: The available data is somehow not in the right form to create the map you want.

Too much irrelevant data?

Data is not spatial?
- has coordinates?
- has addresses?
- has neither, but . . . .
### Data Prep Techniques

**Too much irrelevant data?**

- **query or “select” a subset of the data**

**Data table is not spatial?**

- **has coordinates?**
  - plot XY coordinates
- **has addresses?**
  - geocoding
- **has neither, but . . . .**
  - “join” data table to a matching shapefile

**Spatial data is on the wrong spatial units or you need to combine two spatial datasets?**

- “spatial join” two GIS layers based on location
GIS Data Queries
Why subset? . . . . Your answer will help decide on a method.

- Geographic extent is greater than needed for a small study area.
- Highlight a subregion for cartographic purposes.
- Perform analysis on a subregion.

→ Use simple point and click select.
• Subset based on attributes instead of physical location.
• limit to certain years
• limit to certain populations
• include or remove features with certain characteristics
• eliminate bad values
• Subset based on a combination of attributes.

→ Use “Select by Attribute.”
Choose which values for which field(s) will be the basis for selection.
Calculator interface helps build a SQL expression to query the table.

Compose an expression.

- Logical operators: <, >, =, <=, >=, =, >, <, <=, =, >=, =, >=
- List of available values: 'False', 'True'

Example SQL expression: "KeyRequire" = 'False' AND "PaymentReq" = 'False'
We could pare this subset down even further.

Screened out 572 bathrooms that required a key or payment.
Data selection is ephemeral.

While the selection is active, tools run on the layer will only operate on the selected features.

To make several different selections or to set a symbology for a selection, you can “create layer from selected features.”

To keep a permanent copy of the selected subset for re-use, export it as a shapefile.
Mapping Aspatial Data
Some interesting data are not available as GIS shapefiles or feature classes. The most common format is .csv – a simple text list with values separated by commas. If the file has fields for latitude and longitude (or X and Y coordinates), we can map the records as points.

<table>
<thead>
<tr>
<th>Company Name</th>
<th>Executive First Name</th>
<th>Executive Last Name</th>
<th>Address</th>
<th>City</th>
<th>State</th>
<th>ZIP Code</th>
<th>Record Type</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>CallLiteMiners.csv</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Decide whether the coordinates are lat/lon or something else. Lat/lon values are between -180 and 180. Large numbers indicate meters or feet.
An events layer is also ephemeral unless you export it. Most tools will not run on an events layer.
What if there are no latitudes and longitudes given? Or if we don’t want the information as points?

→ IF we have a geographic layer (vector)

AND

→ IF the values in some field from the non-spatial table exactly match values from a field in the attribute table for the geographic layer

→ THEN we can “join” the two tables.
When joining data, one table stays intact.

This is the left-hand table, called the **target layer**.

The other table gets chopped apart and matched to the target layer.

This is the right-hand table, called the **join layer**.

After the join, records from the target layer get extra columns containing the relevant information from the join table.
The fields that match exactly are called the join **key**.
New fields from joined table, nulls where there was no match.

Joined tables are also ephemeral, and joins are easy to remove. If you will need to use the joined layer repeatedly, be sure to export.

- Note that this join was "one-to-one." The states layer still has 51 records. Where there were duplicates, the first record was chosen. There are ways to join one-to-many and many-to-one.
Spatial Joins
Sometimes it can be helpful to join two spatial layers together, based on the spatial locations of their features.

For example, it might be useful to know which county each mine is in. Or, we might want to know which counties have a mine and how many.
In ArcMap, the join dialog box has a spatial option. There is also a toolbox tool for Spatial Join.

When the points are used as the target layer (left-hand), new fields are added showing which county each falls in.

This join is many-to-one -- several mines are joined to the same county, with duplicate copies of the county record.
When the county polygons are used as the target layer, new choices have to be made.

New fields are added to the counties, including a count of how many cat litter mines each county contained, and most unhelpfully, the average zip code and coordinates of the mine. Be careful when you ask for statistics!!

This join is one-to-many, where the many mines are summarized to one value per county.
The “count” field allows us to symbolize the map to show the abundance of cat litter mines by county.

*Note: in ArcMap, the output of a Spatial Join is durable – a new saved shapefile.*